AD A119307

RETENTION OF SOLDIERING SKILLS:

REVIEW OF RECENT ARI RESEARCH

Andrew M. Rose, Donald H. McLaughlin,
and Daniel B. Felker
American Institutes for Research

and

Joseph D. Hagman
Army Research Institute

TRAINING TECHNICAL AREA

DTIC FILE COPY

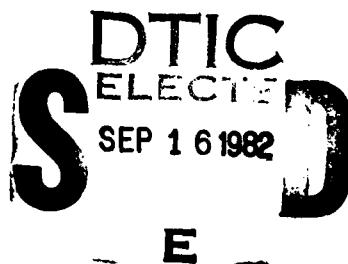DTIC
ELECTE
SEP 1 6 1982
E

U. S. Army
Research Institute for the Behavioral and Social Sciences

March 1981

82 09 16 002

U. S. ARMY RESEARCH INSTITUTE

FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency under the Jurisdiction of the

Deputy Chief of Staff for Personnel

JOSEPH ZEIDNER
Technical Director

L. NEALE COSBY
Colonel, IN
Commander

Research accomplished under contract
for the Department of the Army

American Institutes for Research

NOTICES

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
Technical Report 530	AD-A119307	

4. TITLE (and Subtitle)	5. TYPE OF REPORT & PERIOD COVERED
RETENTION OF SOLDIERING SKILLS: REVIEW OF RECENT ARI RESEARCH	
	6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)	8. CONTRACT OR GRANT NUMBER(s)
Andrew M. Rose, Donald H. McLaughlin, Daniel B. Felker, Joseph D. Hagman	MDA 903-81-C-AA01

9. PERFORMING ORGANIZATION NAME AND ADDRESS	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
American Institutes for Research 1055 Thomas Jefferson Street, N.W. Washington, D.C. 20007	2Q263743A794

11. CONTROLLING OFFICE NAME AND ADDRESS	12. REPORT DATE
U.S. Army Research Institute for the Behavioral and Social Sciences, 5001 Eisenhower Ave., Alexandria, VA 22333	March 1981
	13. NUMBER OF PAGES
	46

14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report)
	Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Skill acquisition	Skill maintenance
Skill retention	Task proficiency
Task acquisition	Skill proficiency
Task retention	

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Sixteen research studies sponsored or conducted by ARI on the acquisition and retention of Army tasks and skills were reviewed. The findings of these studies were synthesized to determine what is known about how soldiers maintain or lose skill proficiency over time.

Some factors associated with training approaches, the nature of soldiering tasks, and individual soldier differences were identified as influencing the retention and loss of skills. The variables investigated in these studies and the experimental procedures used will guide the development and execution of a

DD FORM 1473
1 JAN 73

current research study aimed at developing a practicable method for Army commanders to assess and maintain skill proficiency in units.

Technical Report 530

RETENTION OF SOLDIERING SKILLS:

REVIEW OF RECENT ARI RESEARCH

Andrew M. Rose, Donald H. McLaughlin,
and Daniel B. Felker
American Institutes for Research

and

Joseph D. Hagman
Army Research Institute

Submitted by:
Milton S. Katz, Chief
TRAINING TECHNICAL AREA

Accession For

NTIS GRA&I   ☒
DTIC TAB   ☐
Unannounced   ☐
Justification____

By____
Distribution/
Availability Codes

Dist	Avail and/or Special

Approved by:
E. Ralph Dusek, Director
TRAINING RESEARCH LABORATORY

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES
5001 Eisenhower Avenue, Alexandria, Virginia 22333

Office, Deputy Chief of Staff for Personnel
Department of the Army

March 1981

iii

FOREWORD

The Training Technical Area of the Army Research Institute for the Behavioral and Social Sciences (ARI) has a successful history of using the methodology of experimental psychology for solving Army training problems. Recent research has addressed the problem of how to maximize the retention of military tasks and thereby enhance combat readiness of the fighting force.

Many military tasks acquired during service school training are forgotten over the prolonged intervals of no practice that occur once soldiers are stationed in their units. Although units are responsible for sustaining task performance, opportunities for refresher training are limited because of constraints on time, equipment and other required resources.

This report describes the results of 16 research projects conducted or sponsored by ARI to determine the effects of procedural, task and ability variables which influence attempts to improve task retention through training. These results provide valuable information that will assist Army trainers and training course developers in deciding how, what, and who to train to achieve maximum training effectiveness. When properly applied within the unit environment, this information will help to promote both enhanced task retention and effective allocation of limited training resources.

JOSEPH ZEIDNER
Technical Director

v

RETENTION OF SOLDIERING SKILLS: REVIEW OF RECENT ARI RESEARCH

BRIEF
_____

Requirement:

    To review a set of 16 ARI-conducted and ARI-supported
studies that focus on the acquisition and retention of Army
skills; to identify the variables examined in these studies
that affect learning and forgetting of skills; to synthesize
what is known about performance loss over time for typical
soldiering tasks; to identify training methods and other
procedures that effectively maintain performance skills; to
incorporate findings of the studies into the current research
effort to develop a convenient and practicable method for
Army commanders to assess personnel and job factors that main-
tain proficiency of skills.

Procedure:

    Reports of the 16 research studies were acquired and
analyzed by AIR. These studies encompassed research on about
200 different Army tasks from different MOS ranging from the
combat arms to the combat support services. The purpose,
method, variables investigated, and the results of each study
were described. In several cases, AIR obtained the original
data and conducted secondary analyses to extend the original
findings. A conceptual framework was developed that categor-
ized approaches for dealing with forgetting Army skills and
for describing the synthesis of the results of the 16 research
studies. Research issues and the variables found to affect
skill acquisition and retention were described and organized
within this framework in terms of (1) Training Considerations,
(2) Task Considerations, (3) Individual Soldier Differences,
and (4) Retention Considerations.

Findings:

    As a whole, the research studies were suggestive, but not
conclusive about the factors that influence the learning and
retention of Army skills over time. The most powerful train-
ing factors suggested are the level of proficiency or mastery
that is set as the criterion for ending training, the use of
practice and test trials during training, and the use of
structured training materials. The task factors that appear to
affect forgetting are the difficulty of the task, tasks which
have steps that do not follow from preceding steps, and tasks
which are viewed by soldiers as being less critical. The

individual difference factors that seem to influence how well tasks and skills are remembered are the GT scores and mental categories. How and when retention of tasks is measured will affect the estimates of actual retention and must be considered in any interpretation of skill proficiency.

Utilization of Findings:

The ARI research done on the acquisition and retention of tasks and skills has identified some key variables that will be explored further. This research has high external validity because it examined soldiers learning and forgetting tasks in Army schools and units. The results and procedures of these studies will guide the development and execution of current research on retention of soldiering skills.

RETENTION OF SOLDIERING SKILLS: REVIEW OF RECENT ARI RESEARCH

CONTENTS

LIST OF TABLES

LIST OF FIGURES

RETENTION OF SOLDIERING SKILLS:
REVIEW OF RECENT ARI RESEARCH

INTRODUCTION

A reality that field commanders in the Army constantly face is that all of their soldiers are not able to perform assigned tasks at the desired level of competence. The situation is further complicated by individual differences. Not only do individual soldiers perform some tasks better than others, but different soldiers vary in their proficiency on different tasks. The sum of the proficiencies of individual soldiers at any point in time largely defines a unit's operational capability.

This, of course, is an issue of great concern to the Army. The Army very likely will not know when they will have to mobilize. Time will not be set aside for the Army to upgrade and refine skills of its units before entering a conflict. The next war may well be what has been termed a "come as you are" war. Thus, it is crucial that the Army maintain critical skills at an acceptable level of proficiency at all times.

But these skills are very difficult for the Army to sustain. All skills deteriorate over time. Some are completely forgotten, others only partially so. Tasks vary in how difficult they are to learn and in how quickly they are forgotten. For some tasks the skills needed are quickly recovered with additional training. For other tasks it takes much longer to relearn. And always, the Army has only limited training resources and time for maintaining proficiency.

We should be able to discover through research how well typical soldiers learn to perform a task during training, how quickly they forget the task with and without practice, and how quickly the task can be relearned with different training strategies. This information would be invaluable in the field. Troop commanders, when given a mission to accomplish, could turn to a "task performance book" in which they could look up the kinds of tasks involved in the mission, estimate the current level of troop proficiency on these tasks, and then decide whether the expected level of proficiency is adequate for the mission or whether some type of refresher training is needed.

To generate the kinds of information that ultimately can be used to construct such a convenient and practical guide for field commanders, research will have to be conducted in at least the following areas:

- *The individual abilities of incoming soldiers and the effects of these abilities on learning critical skills.*

- *The level of proficiency attained on critical skills at the end of entry level training.*

1

- The characteristics of different tasks that are trained during entry level training and that are actually performed in the unit.

- The elapsed time interval between training on a task and performing it in the unit.

- The nature of assigned duties in the unit following entry level training.

All of these factors influence the acquisition and retention of skilled task performance. Researchers will have to determine the effects of these factors on retention in order to develop the kind of "task performance book" that would be of practical value to field commanders.

Toward this end, the Army Research Institute (ARI) has undertaken a series of research projects that focus on the acquisition and retention of Army skills. These efforts are especially significant in that they have gone outside the laboratory and examined the learning and forgetting of Army tasks in Army settings. It has been a specific goal of ARI over the last several years to carry out research designed to evaluate the amount of performance loss over time for typical soldiering tasks and to identify training methods and other procedures that are particularly effective in maintaining performance at high levels over long time intervals.

In this paper, we have reviewed a set of 16 ARI-supported and ARI-conducted projects in order to synthesize what is known about the acquisition and retention of relevant skills in the Army, including the implications for Army training procedures and the use of refresher courses. In addition, the actual data from several of these studies were supplied to us. We have conducted several "secondary" analyses of these studies in order to confirm and extend the original findings. Results of these secondary analyses have been incorporated in the text; we have attempted to identify clearly these analyses wherever they are reported, since they may support conclusions that the original authors did not intend.

These sixteen studies also represent the building blocks for a current AIR project with ARI to "develop a convenient and practicable method for Army field personnel to assess personnel and job factors for purposes of maintaining skill proficiency."

The first section of this report presents a conceptual framework for the discussions of the projects to follow. Several issues are raised that are addressed by the specific projects. Following this first section, the results from the projects pertaining to the issues raised are presented. The last section of this report is an annotated bibliography of the projects.

CONCEPTUAL FRAMEWORK

It is theoretically possible to predict and control the rate of forgetting, if sufficient information is known and if the Army is willing to make various types of "investments." The cost -- in training time, in task design, or in recruitment -- will depend upon the method selected.

There are three fundamentally different approaches for dealing with forgetting. These are indicated in Figure 1. First, an increased investment can be made in making the soldier's task easier. For example, automating critical procedures or developing equipment that provides sequencing cues during the task could help to unburden the soldier. In some cases, this can be done for a relatively small investment, such as labeling or providing detailed handbooks. In other cases, it may involve extensive redesign of sophisticated equipment. In view of the critical importance of skill retention to Army readiness, the task simplification or job-aiding approach could be a powerful and cost-effective way to reduce forgetting.

A second approach to dealing with forgetting is to invest in recruitment and retention of soldiers who possess personal characteristics or abilities that reduce the rate or amount of forgetting. General aptitude and memory abilities, as well as specific task skills, are important predictors of forgetting. However, estimates of the contribution of these factors are not available for specific tasks; thus, to make this approach feasible, there must be additional investments, first in research to determine the relationships of abilities to retention, and second to recruit people who better "match" required ability profiles.

The third approach is to invest more time and effort in training. The better a task is learned, the slower will be the decay of performance over time. Tests given at the end of training may indicate that a soldier is proficient at a task, but these tests are poorly suited to indicate how long the soldier will retain his or her proficiency. There is ample evidence that continuing to practice on some tasks past the normal proficiency criteria will retard forgetting of those tasks. ARI research has been designed to discover the limits of improvement in retention that can be obtained by such "overtraining" and to identify the types of tasks for which overtraining is most effective.

Investment in training is particularly costly, however, in that available time for training is limited; extending training time would cut into the time available for operational duty. Soldiers typically learn several dozens of tasks, some simple and others complex, and if each were learned to the extent necessary to eliminate the need for refresher training, training time would be prohibitive. It is important, therefore, to identify those tasks for which retention is critical and to identify methods of training which most effectively use the available time.
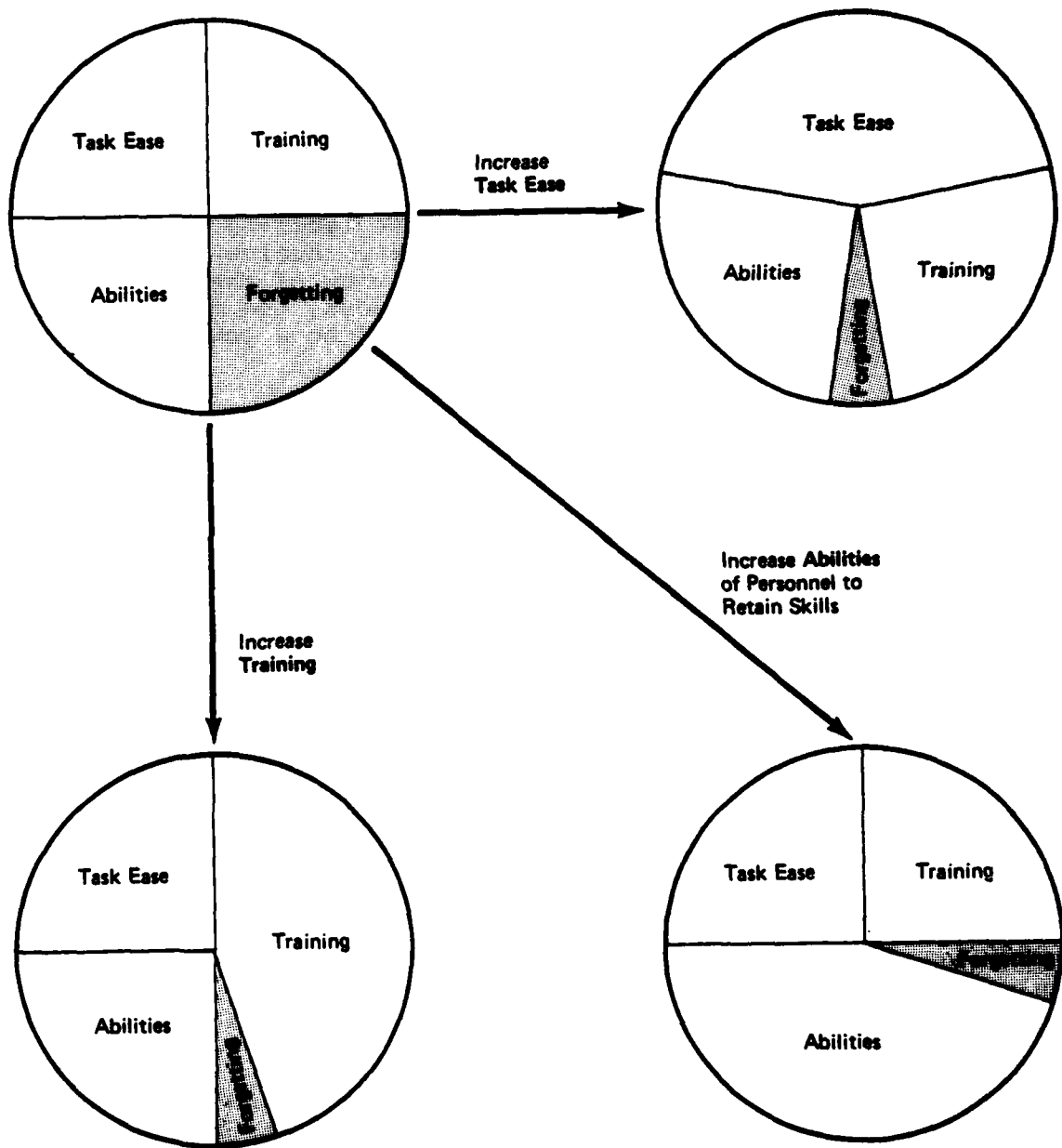
3

FIGURE 1. Alternatives for Reducing Forgetting

4

A similar approach for dealing with forgetting is to compensate for it in the field, either by providing for job practice or by providing refresher training. In a sense, refresher training is one particular version of the training approach, in that it uses the soldier's time for training which would otherwise be spent on other job activities. It is particularly effective in that tests can indicate the soldiers who require refresher training (although motivation to perform on those tests must be adequate); on the other hand, unless refresher training can be successfully accomplished by easily portable training materials, refresher training may require substantially greater investment of resources than original training.

Another aspect of this third approach for dealing with forgetting is to allow it to occur for selected tasks. In some cases this may be appropriate, because the resulting levels of proficiency may still prove adequate. For example, allowing typing speed to decay may be reasonable if the soldier will never be called upon in emergencies to type long documents. For other tasks, this is not an appropriate solution, and forgetting must be delayed.

In summary, the three approaches to retention of skilled performance -- making tasks easier through design, selecting soldiers who possess requisite abilities, and enhancing training, either initially or by providing for refresher training or field practice -- can each affect forgetting. The ARI research examined in this review has investigated various issues related to these approaches.

Table 1 presents an overview of the 16 research reports examined and the issues each addresses. For ease of presentation, we have grouped the issues into four major topics: Training considerations, Task considerations, Individual differences, and Retention considerations. The first three topics correspond to the three approaches outlined above, while the fourth addresses two "technical" aspects of retention studies -- measure variation and the decay function itself. The following sections discuss each topic individually in terms of the relevant research findings regarding specific issues.

FINDINGS

Training Considerations

Training variables have been systematically investigated in nearly a dozen separate ARI efforts in the last several years. The purpose of these efforts has been to determine which methods and conditions of training are associated with particularly stable retention of skills. The dimensions of training variation have included:

REFERENCE	TRAINING CONSIDERATIONS						TASK CONSIDERATIONS				INDIVIDUAL	RETENTION CONSIDERATIONS	
	Mastery vs Proficiency	Presentation vs Test	Spacing	Single vs Sample	Additional Training	Special Methods	Difficulty	Sequential Cue	Memory Requirements	Perceived Relevance	DIFFER-ENCES	Measure Variation	Decay Function
Drussel, 1980						*					*	*	
Goldberg, et al, unpublished	*										*		
Hagman, 1979					*								
Hagman, November 1980		*	*	*					*			*	
Hagman, May 1980	*	*		*					*			*	
Hagman, January 1980		*											
Hagman, in press, 1980		*			*								
Hagman and Schendel, 1979						*							
Mckaupen, et al, 1979											*	*	*
McCluskey, et al, 1978							*	*					
Osborn, et al, 1979							*	*		*			
Schendel and Hagman, 1980	*		*		*								
Shields, Goldberg, et al, 1979							*	*		*			*
Shields, Joyce, et al, 1979													*
Sullivan, et al, 1978							*				*		*
Vineberg, 1975							*				*		

TABLE.1. Overview of ARI Research on Acquisition and Retention

6

- training to "proficiency" or "mastery" criteria
- focusing on presentation or testing activities
- spacing of task repetitions during learning
- focusing on a single task or a domain sample
- additional training (e.g., refresher)
- use of special training materials and methods.

In general, the results of the research can be visualized in terms of a nomograph such as Figure 2. This figure indicates performance loss as a function of time. The straight-line function is a simplification, of course, because forgetting is usually more rapid at first; rate of skill loss typically declines over time. The functions usually observed are negatively accelerated curves. (The straight-line representation is a good approximation to the actual observed curves, if time is measured in logarithmic units, with 1 day, 10 days, and 100 days equally spaced.)

Variation in forgetting rates is portrayed in Figure 3. The four lines on this nomograph may refer to four different tasks, to retention of skill on a single task following four types of training, or to retention of skill by soldiers who differ in their capabilities for learning and retaining information.

Mastery vs. proficiency training. ARI has found ample evidence that significant improvement in retention can be obtained by training beyond the standard proficiency criterion of one perfect performance. Schendel and Hagman (1980) investigated the effect of overtraining on M60 machine gun disassembly/assembly. At the time of original training, some soldiers repeated the procedure (i.e., were overtrained) for a number of trials equal to the number of trials needed to reach the proficiency criterion. When the task was retested eight weeks later, the overtrained soldiers made significantly fewer errors and needed fewer trials to retrain to proficiency. In a separate project concerned with boresighting and zeroing the main gun of an M60A1 tank, Goldberg, Drillings, and Dressel (unpublished) found that requiring three correct performances rather than one (an average of about 20 minutes additional training) significantly improved retention of the procedural steps after five weeks.

In a third project, Hagman (May 1980) investigated the maintenance task of testing alternator and generator output using the 500A Sun Test Stand. Different groups of trainees performed the task 1, 2, 3, or 4 times during training. Results showed that repeating the task more than once significantly reduced performance time and errors on a retention test administered two weeks after training. Furthermore, improvement was directly related to the number of repetitions: the more repetitions, the better the retention. These results are shown in Figure 4.

Thus, in these research efforts, overtraining (in terms of additional trials beyond a proficiency criterion) aided retention. While this basic finding is unambiguous, especially in
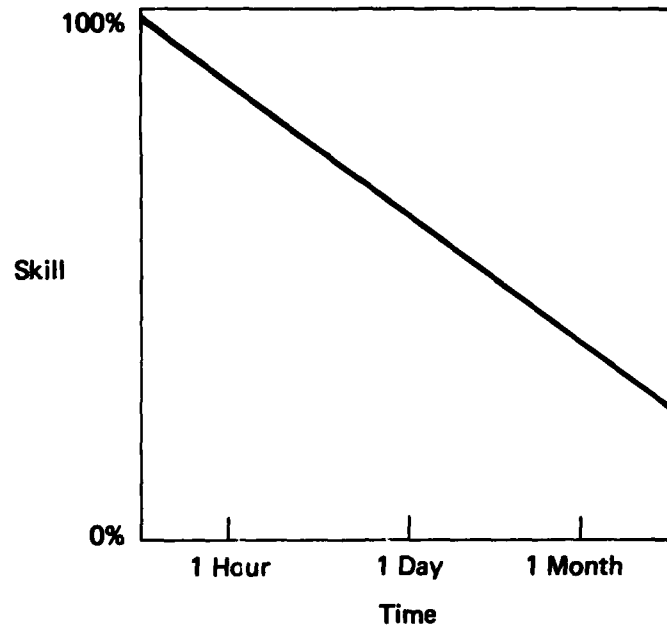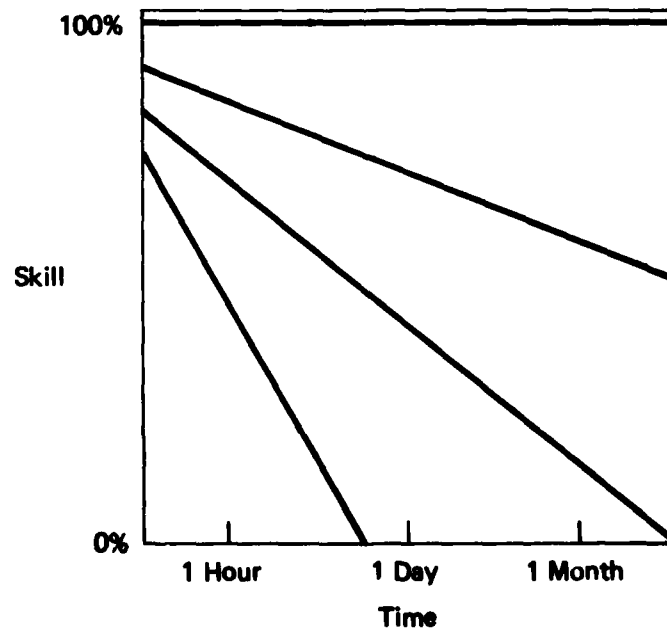
7

FIGURE 2. Hypothetical Retention Function

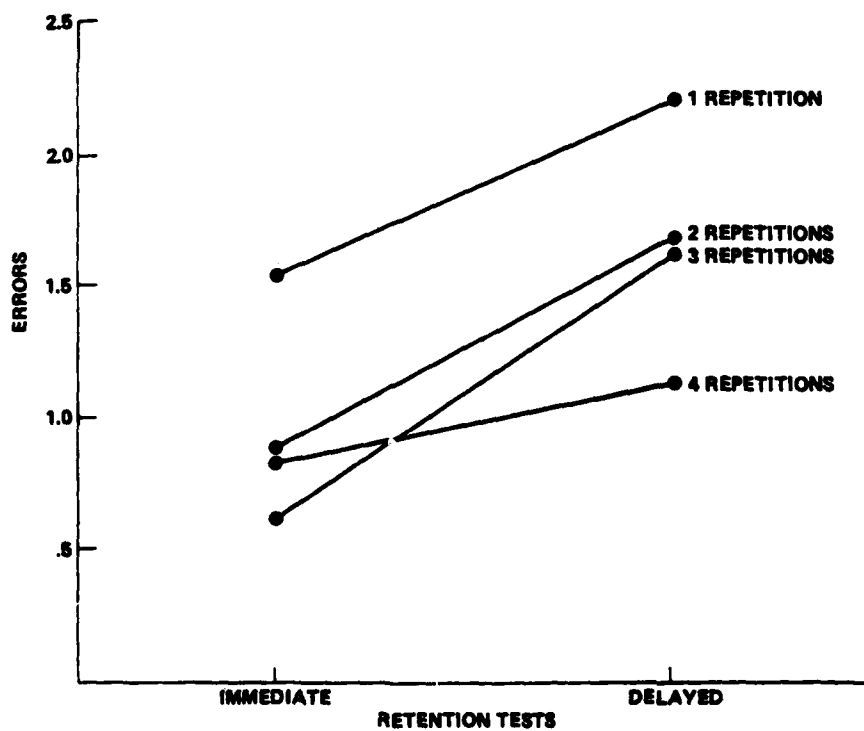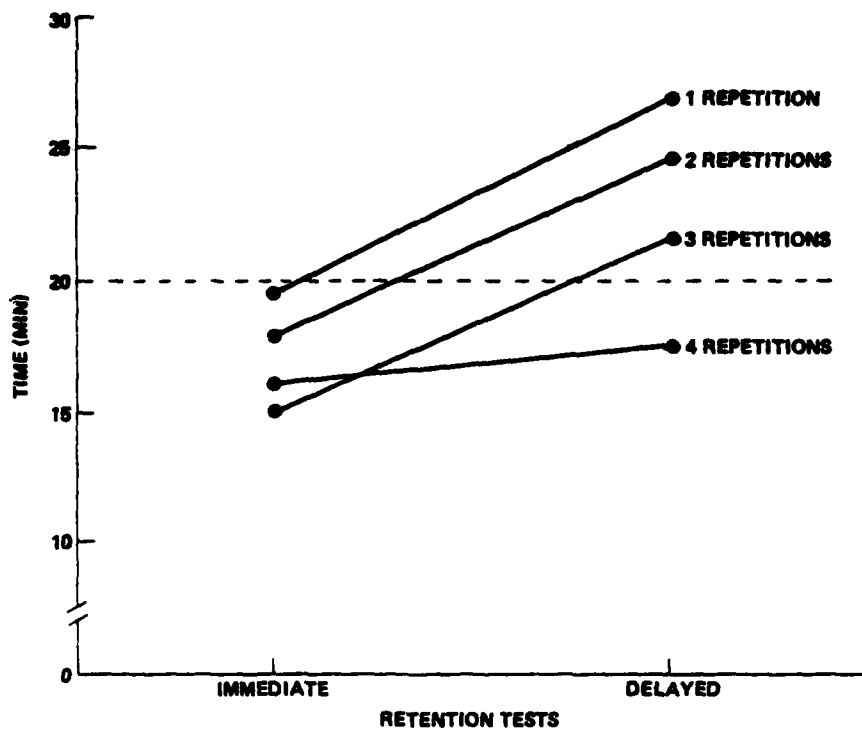**FIGURE 3. Hypothetical Retention Functions**

FIGURE 4. Retention Performance (from Hagman, May 1980).

terms of performance on the first retention trial, an important
practical question remains: How much overtraining need be given?
The answer depends on several factors. The first is the level of
performance required on the initial trial after a no-practice in-
terval. If performance must be at or above a predetermined level
(as is usually the case, for example, for most combat MOS tasks),
the number of original overtraining trials necessary to obtain
this outcome could be determined empirically. The results just
mentioned (Hagman, May 1980) can be used to illustrate this
point. If soldiers must be able to perform the task within 20
minutes after a no-practice interval of two weeks, Figure 4 could
be used to determine how many overtraining trials would be
needed. The horizontal dotted line in the figure shows that
three overtraining trials would be the minimum to reach the
specified level.

On the other hand, if refresher training were a viable op-
tion (e.g., where first-trial performance after a no-practice in-
terval is not critical for mission success), the question of how
much original overtraining to give becomes more complicated. It
becomes a question of resource allocation and cost-benefits: How
can the Army best allocate resources between original training
and retraining to maximize performance effectiveness and minimize
total training costs? The answer would involve estimating or em-
pirically determining the feasibility, relative costs, and rela-
tive benefits associated with various combinations of overtrain-
ing and refresher training trials. While there are cost-benefit
and resource allocation models that are potentially applicable to
this problem, a discussion of them is beyond the scope of this
review.

Without conducting a complex cost-benefit analysis, it is
possible to estimate statistically the point where additional
overlearning trials cease to continue to benefit retention per-
formance significantly. Hagman found that reliable decreases in
task performance time were first found at three overlearning tri-
als, with no additional "benefit" resulting from a fourth trial.
That is, performance on the retention test was significantly bet-
ter for the three-repetition soldiers when compared to the one-
repetition group; furthermore, retention scores for the three-
and four-repetition groups were not statistically different.
Therefore, he concluded that the fourth overlearning trial was
not necessary. Care must be taken, however, in accepting this as
a generalizable conclusion; a different task, a different per-
formance measure, or even a different statistical method for
analyzing the data could result in a different recommendation.

One last issue must be raised before leaving the topic of
mastery training. While it is clear that overtraining aids
retention, Hagman's results raise the possibility of an unwanted
consequence. In addition to the retention test, soldiers in his
experiment were also administered the same task on a slightly
different electrical system. It was observed that performance on
this "transfer" task did not vary as a function of repetition of

10

the 500A Sun Test Stand task. It appears that, in this case, practice on one specific piece of equipment did not aid transfer. In fact, when AIR specifically analyzed these "transfer" data, we found a marginally significant <u>negative</u> effect: there was a slight tendency for errors on the transfer test to <u>increase</u> with number of overtraining trials ($F(1,54)=3.97$, $p<.052$). This result could have been due to a "ceiling" effect: the actual number of errors made was near zero in practically all conditions. Although this "negative transfer" effect was only marginal in this experiment, it is important to consider the possible negative effects on transfer of overlearning a highly similar task.

These efforts have demonstrated the limited value of the training criterion of one perfect performance as an indicator that the soldier will be able to perform the task in the field several weeks later. The rate of forgetting can be substantially reduced (the decay curve can be "flattened") by further initial training. Thus, any guidebook to estimation of retention rates for a variety of tasks must take into account the criterion to which the task was originally trained.

<u>Presentation or practice?</u> An important decision that each instructor and training module developer must make is how to divide the time allocated for training between presentation of information to the trainee and provision of time for the trainee to practice. ARI has explored the magnitude of effects of variation between presentation and practice (or "test") activities.

In two experiments involving a motor skill, Hagman (January 1980 and Hagman, in press) found substantial differences in retention over a one day time span as a function of the allocation of presentation and test trials during acquisition. In both experiments, three groups of subjects were given different sequences of presentation (P) and test (T) trials as follows:

						ACQUISITION													RETENTION
						TRIALS													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	24 hr
Standard	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	P	T	T
Test	P	T	T	T	T	T	P	T	T	T	T	T	P	T	T	T	T	T	T
Practice	P	P	P	P	P	T	P	P	P	P	P	T	P	P	P	P	P	T	T

The first experiment (Hagman, January 1980) involved positioning a bar to estimate a distance of 250 mm. The average absolute error of the estimation after 24 hours was smallest for the "Test" group, intermediate for the "Standard" group, and greatest for the "Practice" group. These different retention losses occurred despite almost equal acquisition among the groups. The advantage

11

gained by emphasizing test trials was replicated in a nearly
identical experiment (Hagman, in press) that presented the task
as a position location task instead of a distance estimation
task. In this experiment, the average absolute errors after 24
hours were again smallest for the "Test" group, intermediate for
the "Standard" group, and greatest for the "Practice" group.
(However, in this experiment, differences were also noted during
acquisition.)

A third experiment, by Hagman and Schei...el (1979), focused
on a practical task of formatting and typing military correspon-
dence. This task is customarily performed with formatting direc-
tions available to the typist. The authors compared the outcomes
of refresher training when directions were available at all times
with outcomes when the directions were not available. The out-
come measures, average total amount of time taken to complete all
tasks and total number of errors, were recorded with the direc-
tions available. Although the results of this experiment are
somewhat ambiguous because of pre-existing differences between
the two treatment groups, the results suggest that there are no
effects of withholding the directions during two refresher ac-
quisition trials.

Finally, in an examination of Training Extension Courses
(TECs), Holmgren et al. (1979) investigated five combat tasks and
found that retention over a two-month period was slightly better
for groups of Active Component (AC) and National Guard (NG) mem-
bers when testing accompanied the TEC instructional module. The
effect was small but was of the same order as the difference
favoring TEC training over conventional training (see below).

These experiments taken together indicate that practice, or
testing, is very important for retention. Reducing the amount of
time in which the trainee has access to training information
(presentation trials) is not critical, as long as there is at
least some presentation. The results obtained by Hagman
(January, 1980 and in press) are particularly important in that
they indicate a positive effect of practice even when no feedback
is given. The results of those two experiments might be ex-
plained in other terms, however; namely, that the groups for
which every test trial was immediately preceded by a presentation
trial were learning a different, short-term memory task and were
ill-prepared to recall after an extended period.

Spacing of learning trials. There is a substantial litera-
ture that suggests that allowing a noticeable interval between
repetitions of a task to be learned increases the retention of
performance skills. Because of the scheduling problems that
would be created by planning spaced training programs, it is im-
portant to know the extent of this phenomenon. Is it substantial
enough to warrant the extra investment required? In order to
demonstrate the applicability of the spacing phenomenon to Army
tasks and to explore the limits of the effects of distributing
practice, ARI has undertaken two research projects focused on the

12

distinction between massed and spaced task repetitions during learning.

In an experiment involving 60 trainees, Hagman (November 1980) investigated the effects on retention of one-day spacing of repetitions of a task involving testing alternator and generator output using a 500A Sun Test Stand. He found that three trials spaced a day apart were dramatically more effective than three trials without intervening activities. This was found in spite of the fact that soldiers in the "spaced" condition also learned other tasks during the time between repetitions. After an interval of 14 days, the soldiers in the "massed" condition took 50% longer to do the task and made over twice as many errors as they did at completion of training. Soldiers in the "spaced" condition showed only a 6% increase in time and a 15% increase in errors.

On the other hand, in an experiment involving disassembly/assembly of the M60 rifle by 38 reservists, Schendel and Hagman (1980) found little effect of "spacing". In this study, soldiers were trained to a proficiency criterion and then given an equal number of extra trials, either immediately or after four weeks. Eight weeks after the original training, there were essentially no differences between these groups, although both groups performed substantially better than a control group that received no extra trials. In this situation, spacing of training sessions appeared to be unimportant.

Examination of the differences between these two experiments suggests possible limits on the spacing effect. Of the many differences, one that plausibly accounts for the different findings is the variation in acquisition criteria. In the first experiment, it seems likely that very few soldiers learned the task really well: only three task repetitions were given. In the second experiment, however, both groups learned far beyond a simple proficiency criterion. Another plausible difference is the distinction between "sessions" and "trials." In the Hagman experiment, individual trials were spaced or massed, while in the Schendel and Hagman experiment, the training sessions (i.e., either training to proficiency or overtraining) were spaced or massed.

Single task or domain sample? Many soldiering tasks fall into clusters. During training a decision must often be made of which specific tasks in a cluster to focus on and how many different tasks to include. With limited time, a choice is usually made between learning one task very well or several tasks less well. Alternatively, however, it might be possible to select a sample from a domain of tasks and present them in such a way that they maximize learning and retention on all the tasks in the domain.

This was explored in the experiment by Hagman (November, 1980) mentioned above. In that experiment, half of sixty

13

students were trained on the task of testing charging system output using the 500A Sun Test Stand, and the other half were trained by presenting the same task on three similar but different charging systems. Retention fourteen days later was measured both on the original and on a fourth charging system. When performance was measured on the original device, on which one group had three acquisition trials but the other only one, the performance of the two groups was identical. (This was probably due to a "ceiling" effect; both groups made very few errors.) When measured on the fourth, "transfer" system, no significant differences were found. Thus, this experiment raises hopes that training can be delivered on a sample of tasks without reducing the learning of any one of them.

There are some obvious cautions, however. For example, in a separate experiment using the same charging system task, Hagman (May, 1980) found that increasing repetitions of the task (on a single device) increased retention of that task. However, there was also a tendency for errors on a test of performance on a related device to increase. Thus, by focusing on a single task (or device) in the domain, the training may induce task-specific responses that interfere with performance on related tasks.

In order to decide which tasks to include in a training module, it will be necessary first to define the scope of the domain. In many cases, a single task is sufficiently different and important that training should focus on it alone. But there are many processes that contribute to performance on a wide variety of tasks. These are processes such as paying attention to safety precautions, using guidebooks to look up procedures, and using strategies for stable acquisition of skills. The Army's new Basic Skills Education Program may provide a vehicle for teaching these general processes, so that soldiers are prepared to generalize from instruction on a single representative of a task domain.

Additional training. Retention is a function of the time since the task has been practiced, so retention will be better for soldiers who have undergone refresher training or who, as part of their regular duties or for other reasons, have practiced the task in question. The effects of refresher training were discussed in summarizing the training factor of spacing task repetitions. An interesting finding was that in one case extra training during initial acquisition was as effective as refresher training after four weeks, in terms of retention after eight weeks (Schendel & Hagman,1980). In another experiment, Shields, Joyce, and VanWert (1979) found effects of a refresher course on retention, but on only one out of six tasks was the effect statistically significant. However, using more powerful statistical tests would likely show significant effects on other tasks.

Uncontrolled practice has confounded the results of some of the experiments, although most experiments have eliminated soldiers who report having practiced. More extensive information on the effects of practice is needed.

14

Special methods. There is a need to refine training methods so that greater retention of skills can be achieved for the same investment of training time. ARI has carried out three efforts since 1978 that evaluate particular innovative training techniques, with varying success.

First, Holmgren et al. (1979) studied the effectiveness of Training Extension Courses (TEC) on retention of skills in four combat areas and on common skills. In this effort, they taught two groups of active soldiers (AC) and of National Guardsmen (NG) using the self-paced audio-visual TEC modules, and two groups using conventional instruction. (The two groups in each case differed in the inclusion of tests before and after training.) Performance was measured in terms of number of steps passed, both immediately after training and eight weeks later; and in almost every case the results favored the TEC procedures. The authors of this study attempted to isolate the proportion of steps learned from the proportion of steps that could be guessed without training, by comparing performance to an untrained control group. Furthermore, they estimated retention as a ratio of the proportion of learned items retained after eight weeks to the proportion of learned items immediately after training. The results of their estimation indicated that although the soldiers who received TEC training performed more accurately, their proportional loss over eight weeks was comparable to that of soldiers receiving conventional instruction.

Figure 5 shows two hypothetical effects of improvements in training. The top portion shows the effect on the retention function when there is a "true" increase in the resistance of skill to forgetting: the slope of the function "flattens out." The bottom portion of the figure shows another possible effect: skill level is increased, but the retention function remains parallel to the original line, and there is no increase in resistance to forgetting. In terms of the nomographs in Fig. 4, the effect of the TECs was to raise the retention function but not to flatten it.

Another experiment, by Dressel (1980), achieved neutral results in an evaluation of mnemonic enhancement of learning. Based on numerous studies that demonstrate positive effects of verbal elaboration on memory for sequential information, this experiment examined the effects of introducing verbal strings, or acrostics, as cues for sequential steps in the placement of antipersonnel mines. Comparison of two groups introduced to the mnemonic strategy with a control group (given standard instruction) on one-month retention favored neither group. This was probably due either to a "ceiling" effect (i.e., the task was quite simple to start with) or to the fact that that most soldiers in the mnemonics groups failed to use the strategy. Those few who did use the strategy outperformed the others. The conclusion one is drawn to is that more concerted efforts should be made to convince trainees of the power of verbal elaboration prior to beginning the acquisition of critical information.

Skill

Time
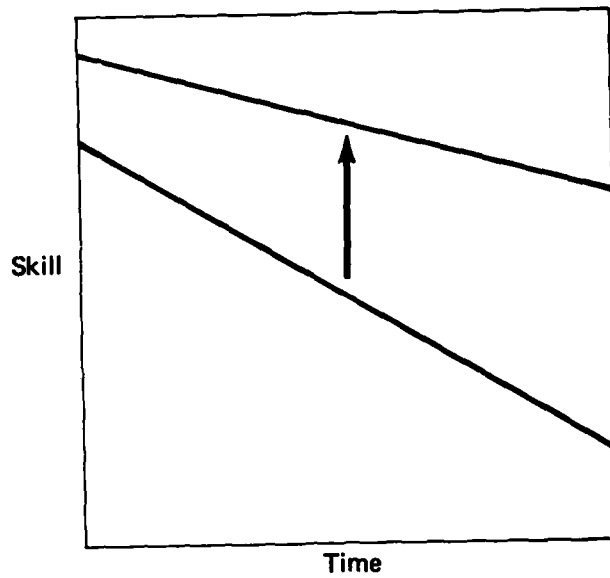
Increasing Resistance
of Skill to Forgetting

Skill

Time
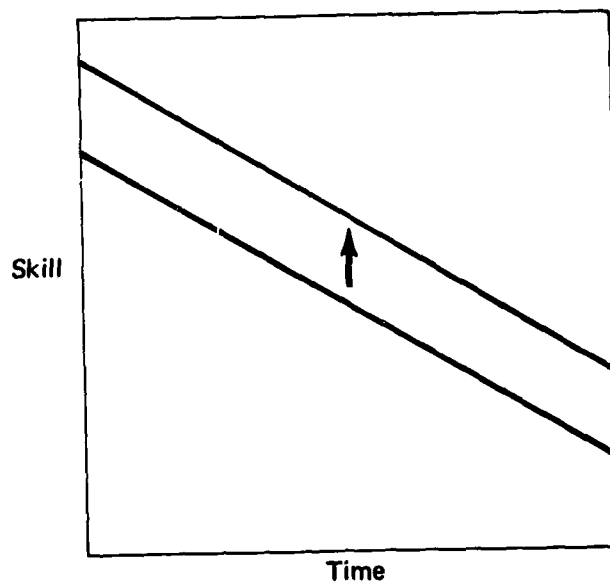
Increasing
Skill Level Only

FIGURE 5. Two Effects of Improvements in Training

Finally, Sullivan et al. (1978) undertook an evaluation of the use of alternative training methods for individuals with different aptitude profiles (that is, an "Aptitude-Treatment Interaction" or ATI study). This effort produced ambiguous results with respect to retention, however, in that for most performance measures the retention scores were higher (although insignificantly) than the immediate post-training scores. It is not clear whether these results were due to learning which occurred in the one-month retention interval or merely to the very small numbers of subjects available for retention testing.

Summary. Overall, the most powerful training factor appears to be the level of proficiency or mastery that is set as the criterion for ending training. Clearly, much improvement in retention can be obtained by going beyond the criterion of one perfect performance. Secondly, it is very important to include practice, or test trials, even at the cost of spending less time on actual presentation of information. Third, although there are advantages of spaced versus massed repetitions during learning, the effect is limited. Fourth, although the results are not conclusive, gains can be obtained in acquiring and retaining domains of tasks by sampling more than one task from the domain during training. Finally, the TEC method for training appears to produce higher acquisition and retention levels than conventional training for some tasks.

Differences in Retention Among Tasks

If information were available on what tasks or what types of tasks are most susceptible to forgetting, steps could be taken to reduce that forgetting through refresher training, for example. Shields, Goldberg, and Dressel (1979) proposed the development of nomographs indicating for each task its associated decay rate, and in carrying out its research program ARI has selected a broadly representative set of soldiering tasks. The variety of tasks studied in the various efforts is indicated in Table 2. Values for decay rates on these tasks can be computed; however, a comparison of these rates is questionable because of the multitude of sources of variation, such as the effectiveness of training, the types of individuals being tested, the types of measures used to estimate decay, and design constraints on the studies (e.g., presence or absence of uncontrolled practice during the retention interval).

We will examine task considerations from two perspectives. First, we will discuss specific task (or step) features addressed in the ARI research. Second, we will take a slightly broader view of possible task dimensions or characteristics which could be used in the development of a general "task taxonomy" suitable for classifying tasks on the basis of predicted retention loss.

TABLE 2.

Army Tasks Studied in Recent ARI-Supported Acquisition and Retention Research

Dressel, 1980

Install M14 antipersonnel mine

Goldberg, Drillings, and Dressel, unpublished

Boresight the M60A1 main gun
Zero the M60 A1 main gun

Hagman, June, 1979

Straight–copy typewriting

Hagman, May, 1980 and November, 1980

Testing charging system output using the 500A Sun Test Stand generator:	Base settings
" "	Cable connections
" "	Switch positions
" "	Testing procedures
" "	Shutdown

Hagman, January, 1980, and in press

Simple horizontal motor movement

Hagman and Schendel, October, 1979

Straight-copy typewriting
Job-task typewriting

Holmgren, et al., April, 1979

M60 MACHINEGUN:

Clear
Disassemble
Inspect
List defects
Assemble
Perform functions checks
Identify grip position
Demonstrate assault position
List rates of fire

M551:

Identify center of target mass
Distinguish reticle alignment
Align reticle
Determine range

TADDS:

Select site
Deploy ground planes
Perform march order
Perform self-test check
Perform voice check
Orient TADDS
Respond to alarm
Charge battery
Replace battery

SQUAD RADIO:

Identify parts
Plan battery life
Identify planning range
Perform oper. maintenance
Assemble
Operate
Identify jamming signals
Recognize malfunctions

TUBE ARTILLERY:

Tune and install
Use quadrants for high-angle fire
Perform micrometer test
Perform test for positive and negative correction

Tune and install
Energize
Perform battery voltage check
Prepare for radar check
Perform radar check
Plot FAAR on display
Perform "plot to tell"
Perform preventive maintenance

TABLE 2.(Continued)

McCluskey, et al., November, 1978

Select temporary battlefield positions
Use challenge and password
Estimate range
Perform operational maintenance on field telephone
Describe engagement of armored vehicle by individual and crew weapons
Operate M203 grenade launcher
Identify enemy vehicles
Apply four life-saving measures
Operate M60 machinegun
Identify NBC hazards
Install/recover/fire Claymore mine
Apply first aid
Collect/report information: SALUTE
Maintain M203 launcher and ammunition
Process enemy personnel
Prepare and restore M72A2 LAW

Schendel and Hagman, July, 1980

Disassemble and assemble M60 machinegun

Osborn, et al., December, 1979

DRIVER TASKS:

Check M27 periscope
Remove M27 periscope
Check M24 periscope
Install M24 periscope
Operate M24 periscope
Check gas particulate unit
Prepare to fire
Check track tension

GUNNER TASKS:

Check gas particulate unit
Charge manual elevation system
Prepare telescope
Prepare periscope
Place turret in operation
Prepare-to-fire (A)
Prepare-to-fire (B)
Prepare-to-fire (C)
Prepare to boresight
Prepare azimuth indicator
Operate elevation quadrant
Boresight telescope
Boresight periscope
Boresight IR sight
Boresight searchlight
Boresight coax

COMMON TASKS:

Read replenisher tape
Know target acquisition procedures
Load .45 pistol
Clear .45 pistol
Disassemble .45 pistol
Assemble .45 pistol
Clear .45 stoppage

LOADER TASKS:

Check engine transmission oil
Check track tension
Adjust track tension
Check gas particulate unit
Stow main gun rounds
Load Main gun (A)
Load Main gun (B)
Load Main gun (C)
Ready coax
Clear main gun misfire
Unload main gun misfire
Reduce coax stoppage
Clear coax
Remove coax
Disassemble coax
Assemble coax
Check coax operation
Mount coax
Load coax
Remove breech block
Disassemble breech block
Assemble breech block
Install breech block
Prepare for boresight
Boresight coax
Check main gun alignment
Prepare-to-fire

Load submachinegun
Clear submachinegun
Disassemble submachinegun
Assemble submachinegun
Clear submachinegun stoppage
Give first aid to burn
Give first aid to fracture

19

TABLE 2. (Continued)

Shields, et al., September, 1979

Cardiopulmonary resuscitation
Stop bleeding
Challenge and Password: One man approaches
Challenge and Password: Group approaches
Report enemy information: Size/Activity/Location
Report enemy information: Unit/Time/Equipment
Don protective mask
Individual defensive position: Outline
Individual defensive position: Describe
M60 machinegun - Load/fire
M60 machinegun - Reduce stoppage
M60 machinegun - Unload/clear
M203 grenade launcher - Disassemble/assemble
M203 grenade launcher - Load/fire
M203 grenade launcher - Reduce stoppage
M203 grenade launcher - Clear
M72 Light Anti-tank Weapon (LAW) - inspect/fire
M72 LAW -Restore
M16 rifle - Disassemble/assemble
Communications check

Shields, et al., March, 1979

Pre-energize the M54 launch station
Energize M54 launch station
De-energize M54 launch station
Before-operations PM checks on the M730
Install and operate the TA-312/PT telephone set
Emplace, check, and adjust target alert data display set AN/GSQ-137

Vineberg, June, 1975

Drill and Ceremonies (D&C)
  Marching movements for the individual
  Manual of arms executed from sling arms
First Aid (FA)
  Apply tourniquet, dress wound, treat for shock
  Treat for burns, treat for shock
Individual Tactical Training (ITT)
  Passage of obstacles during daylight
  Individual maneuver techniques
Guard Duty (GD)
  Inspecting officer
  Hour of darkness with proper or improper authorization
M16A1 Rifle (M16)
  Clearing the weapon
  Immediate action
Chemical, Biological, and Radiological (CBR)
  First aid for a nerve agent casualty
  Reaction to a nuclear burst without warning
M60 Machine Gun (M60)
  Placing the M60 machinegun into operation and performing immediate action

Several efforts carried out by ARI have included post hoc descriptive analyses of the types of task steps that are most frequently forgotten. Four types of task or step variations have been identified as predictive of performance on retention tasks: (1) difficulty or skill level required, (2) absence of sequential cues for task components, (3) memory requirements (absence of printed guides), and (4) perceived relevance of the steps to the goal of the task.

Difficulty. In exploring retention on sixteen common soldiering tasks, McCluskey et al. (1978) found lower performance on tasks rated "difficult" than on those rated "easy." This was true for both training posttest and retention tests given three, seven, and eleven weeks later, but the difference remained relatively constant, as did the overall performance level. Practice apparently occurred on the tasks studied in this project. In an investigation of tank crewman skills, Osborn et al. (1979) identified the step on each task requiring the highest skill level and found these steps to be among the most frequently forgotten.

In a more straightforward investigation of task difficulty, Vineberg (1975) examined performance on thirteen subtests of the Comprehensive Performance Test (CPT) after finishing BCT and six weeks later during AIT. He found that the various CPT subtasks varied considerably in difficulty, as measured by baseline performance. Furthermore,

> ...the average level of performance on individual
> subtests during retention testing (was) related to
> the average level of performance on these tests
> during baseline testing. That is, subtests tend[ed]
> to maintain their ordering with respect to the prob-
> ability that they will be passed.   (p. 15)

The findings from McCluskey et al. are somewhat weakened by the high attrition rate of soldiers involved in the project. If those who were available for retention testing were different from those not available, then the generality of the results would be limited. As a suggestion for future studies, an additional analysis, comparing these two groups on available performance measures, would help to resolve this issue. Another reason for carrying out this additional analysis is to provide empirical evidence needed to generalize from future efforts in which attrition occurs. Adjustments for attrition can then be made in estimating retention rates. For example, if (1) those available for retention testing have relatively higher acquisition scores and (2) higher acquisition scores are associated with slower forgetting, then estimated forgetting rates based on the study would actually be lower (i.e., retention would be better) than what might be expected for the entire Army population.

Sequential cues. Both Shields, Goldberg, and Dressel (1979) and McCluskey et al. (1978) distinguished between tasks in which there was greater or less sequential information leading the

21

soldier from one step to the next. In both cases, forgetting was greater on tasks with steps that did not follow from preceding steps.

Memory requirements. The Army recognizes the value of memory aids and provides guidebooks for tasks in which real-time reference to a guidebook is feasible. Even with guidebooks, however, substantial forgetting occurs. Hagman (May, 1980 and November, 1980), for example, found many errors on a maintenance task in which a guidebook was constantly available. He noted, however, that the majority of errors occurred not in remembering to execute a step but in remembering how to carry out one par-ticularly difficult and ill-described step. At this step, the soldiers were required to retrieve a complex procedure from memory.

Perceived relevance. Individual soldiers evaluate task steps during training and consider some of them critical to per-formance and others to be largely irrelevant to the goal of the task. Steps not considered relevant and frequently omitted on retention tests include safety steps (Shields, Goldberg, and Dressel, 1979; Osborn et al., 1979) and task beginning and ending steps (Osborn et al., 1979).

These results on individual task characteristics are largely exploratory and do not relate to any common taxonomy of tasks. There has been no direct investigation to date of the relation-ships between types of tasks (e.g., such as MOS clusters or psy-chomotor vs. perceptual vs. organizational) and forgetting. Forgetting occurs on tasks in all sectors of the Army, and a field commander can estimate forgetting rates primarily only if he knows the level of training received, the abilities of the soldiers, the retention interval, and the difficulty of the task.

Taxonomic considerations. To restate the issue of task dimensions, we can ask the following question: What do we know about the dimensions or characteristics of tasks that will help us to predict rates of proficiency loss? That is, is there any-thing we can tell about the characteristics or nature of a task that will help us to predict how rapidly the skills needed to perform that task will deteriorate and to know when we can best insert refresher training to build the skills back up to the desired levels of proficiency?

In the aggregate, the ARI research efforts represent a fair-ly solid beginning for understanding the nature, diversities, and complexities of learning and retaining task skills. The in-dividual efforts in one way or another examined approximately 200 different tasks, each in turn consisting of varying numbers of subtasks and steps. These tasks were drawn from several dif-ferent MOS, ranging from the Combat Arms (11B, 19E) to combat support services (71L, 63G). Simply put, tasks vary in how dif-ficult they are to learn and in how quickly they are forgotten. It is clearly impossible to study the specific factors which aid

22

retention of each task and to design training tailored to that particular task. What is needed is to identify a manageable number of task factors which will predict rates of proficiency loss for large numbers of tasks.

One of the major products of AIR's current research project is to develop a task taxonomy that will generate these predictions. The ARI studies done by Shields, Goldberg, and Dressel (1979), Hagman and Schendel (1979), McClusky, Hiller, Bloom, and Whitmarsh (1978), Sullivan, Casey, and Hebein (1978), and Osborn, Campbell, and Harris (1979), all provide clues of some task factors that might be organized into a predictive taxonomy. While the evidence from these efforts is not compelling and we cannot generalize the findings from one study to another, it is instructive to look at the task factors found by these researchers to affect retention.

Shields, Goldberg, and Dressel (1979) tested soldiers' performance on twenty basic common tasks to identify task factors which influence the rate of skill decay in performing these tasks. The tasks evaluated included the assembly/disassembly of the M203 grenade launcher, putting on a gas mask, and reporting enemy information. Soldiers at Ft. Sill who were completing entry-level training and who had completed training during the previous twelve months were tested on the tasks. The researchers found that the factors that accounted for most of the differences in retention were the number of steps in a task, when the task was originally learned, and the presence of subtasks. Thus, simpler tasks and those learned earlier and practiced are retained better. Whether these findings can be replicated and hold for other tasks remains to be studied.

Osborn, Campbell, and Harris (1979) examined the extent to which armor crewmen retain skills between completing basic armor training and the early months of unit performance. Gunners and loaders (19E) and drivers (19F) were tested on a set of common military tasks (e.g., loading/cleaning .45 pistol) and tasks specific to their MOS (e.g., place turret in operation for 19E; check track tension for 19F) at the end of MOS training and again four to eight months later. Proficiency on job-specific tasks increased while performance on common tasks declined at the second testing. This suggests that job specificity may be a task factor that predicts rate of proficiency loss.

Hagman and Schendel (1979) studied a similar task factor but in a completely different context and MOS. These researchers investigated the effects of different training techniques on typing performance. Administrative specialists (71L) were trained and tested on straight-copy typing and job-task typing (preparing military correspondence, memos, etc.). The training improved job-task typing but not straight-copy typing. This finding and the results of Osborn et al. (1979) are consistent in showing that job-specificity of tasks is a factor that appears related to retention and proficiency.

23

McCluskey, Hiller, Bloom, and Whitmarsh (1978) set out to develop skill decay curves for sixteen common tasks for MOS 11B and 11C. The tasks were categorized on the basis of how hard they were to learn and on the amount of performance guidance cues available which serve as prompts or feedback. Samples of soldiers were trained on the tasks and tested immediately and retested a number of weeks later. The study failed to yield reliable skill decay curves due to the unavailability of soldiers for retesting. Nevertheless the two task factors used represent possibilities for further study.

Another way of classifying task dimensions is suggested by the work of Sullivan, Casey, and Hebein (1978). This research focused on the effects of different instructional strategies on the acquisition and retention of different kinds of skills by soldiers with different aptitudes. The skills to be learned were related to employing and engaging Redeye missiles. Of interest to us is that specific tasks differed in the extent to which they involved perceptual, psychomotor, or cognitive demands. These task dimensions have been used by others in creating task taxonomies in non-Army contexts.*

Table 3 summarizes the task factors dealt with in a subset of six ARI-supported efforts. These factors can be used to differentiate tasks along some dimensions that might be related to retention. Ideally, these factors would predict rates of proficiency decay so that a commander, knowing that his troops had to perform "psychomotor" tasks or tasks that had few "performance guidance cues," would know what level of performance he could expect of his unit and what kind of refresher training might be in order.

The task factors surfaced by the ARI research may or may not ultimately prove useful. The results are not conclusive. The questions raised, the approaches attempted, and the constraints uncovered are valid. We intend to build and expand from this start.

Individual Differences

If variations in forgetting could be predicted for different individuals, then methods, for dealing with forgetting could be more efficiently implemented. Soldiers least likely to forget

---

*For example, see J. A. Aagard and R. Braby, Learning guidelines and algorithms for types of training objectives. Orlando, Florida: Training Analysis and Evaluation Group, Technical Report No. 23, March 1976; and A. Mirabella and G. R. Wheaton, Effects of task index variations on transfer of training criteria. Technical Report:NAVTRAEQUIPCEN 72-C-0126-1, Naval Training Equipment Center, Orlando, Florida, August 1973.

TABLE 3. Task Factors Examined by ARI - Supported Research

ARI Reference	Task Factors
Shields, Goldberg, and Dressel (1979)	● Number of task steps ● Order of original training of task ● Subtask structure
Hagman and Schendel (1979)	● Job-specific task
McClusky, Hiller, Bloom, and Whitmarsh (1978)	● Difficulty to learn ● Amount of performance guidance cues
Osborn, Campbell, and Harris (1979)	● Job-specific task ● Common task
Sullivan, Casey, and Hebein (1978)	● Perceptual ● Psychomotor ● Cognitive
Vineberg (1975)	● Difficulty (Baseline performance)

could be selected for critical tasks, training could be prolonged for those who need it, or refresher courses could be targeted to those soldiers most likely to have forgotten a task. The use of a training criterion of a particular number of correct responses recognizes this variation and makes the implicit assumption that requiring achievement of a particular initial learning criterion will remove most of the individual variation in subsequent retention.

Five ARI efforts have addressed the question of predicting retention from individual ability measures, primarily GT scores. In an investigation of use of a mnemonic strategy for remembering an equipment installation procedure, Dressel (1980) found that soldiers with high GT scores retained more of the task than others in the groups assigned to use the mnemonic strategy, but that there was no relation of GT scores to retention in the control group. The most reasonable explanations of this finding are (1) that use of a mnemonic strategy requires a high GT score, (2) that soldiers with high GT scores can more easily adapt to any learning situations requiring a new learning strategy, or (3) soldiers with high GT scores, unlike others, have used mnemonic strategies in the past.

In their project on the effects of TECs on retention of combat skills, Holmgren et al. (1979) tested the earlier finding that TECs tended to compensate for low GT scores, in that post-training performance is less related to GT scores after TEC training than after coventional training. Their results were somewhat disappointing, in that positive correlations occurred as frequently in the TEC treatment as in the conventional treatment.

Both general and specific skill levels were taken into account by Goldberg et al. (unpublished) in their research on tank gun boresighting and zeroing. First, there were no significant relations of performance to GT scores. On the other hand, even though the training criterion was controlled, soldiers who had previous experience as gunners performed significantly better than others.

A complex project to investigate the potential gains to be realized from the aptitude-treatment interaction phenomena ("some people learn better one way and others another") was conducted by Sullivan et al. (1978). Unfortunately, although several aptitude types were identified, no significant relations were found between training method, aptitude, and retention.

Finally, Vineberg (1975) examined performance on thirteen subtasks of the CPT (see above) as a function of Mental Category. He found a direct relationship between Mental Category and test performance both during baseline testing and during retention testing. Overall, the performance of soldiers in Mental Category II was significantly superior to that of soldiers in Mental Categories III and IV. However, there was substantial overlap of scores on all tests; also, the pattern of results was not

identical for all subtests. That is, on some tests, soldiers in
Categories III and IV were actually better than soldiers in
Category II.

The findings concerning individual differences in retention
are far from conclusive, although there is a general consensus
that effects do exist and can be found. In future efforts, to
make substantial headway, a more thorough approach to the
measurement of multidimensional differences in individuals is
necessary. Similarly, an important question for optimizing duty
assignment is whether forgetting rate is a general trait of an
individual or whether different individuals remember different
tasks better. Moreover, if forgetting is a general trait, a
single predictive measure could be used for individual refresher
training on many tasks.

Retention Considerations

Two important questions concerning retention remain and have
been addressed by ARI research. First, how does forgetting af-
fect performance? ARI efforts on retention have reported a
variety of measures, and it is useful both for interpreting
research results and for adapting refresher training to know how
these measures are related. Second, what is the form of the for-
getting function? The straight-line approximations used in
several studies may conceal actual variations in forgetting
rather than reveal them. Because, as is well known, the rate of
forgetting slows as time passes, studies that focus on long in-
tervals may produce lower decay rate estimates than studies that
focus on short intervals.

Measure variation. There are actually two kinds of measures
that affect estimation of retention: real differences and statis-
tical artifacts. Two types of real differences were studied in
the retention projects: speed versus accuracy, and performance on
a single recall test versus time taken to relearn to a specified
criterion. Except on tasks that were extremely speeded, such as
typing, the research found that factors that increased speed in
retention also decreased errors (Dressel, 1980; Hagman, May and
November, 1980) and factors that improved single trial recall
also improved relearning rates (Schendel and Hagman, 1980;
Hagman, May, 1980).

Holmgren et al. (1979) were particularly careful to avoid
statistical artifacts in their measures in that, before estimat-
ing differences between treatment groups, they used a control
group to estimate how much of the task would be accomplished
without error even without training. Two of the efforts used
linear regression to estimate the average rate of performance
loss, and care must be taken to avoid interpreting these results
as indicators that the loss is actually linear.

27

Another "measure variation" issue concerns the choice of statistical methods to use in analyzing results from retention experiments. In a typical experiment, the primary concern is the effect of some treatment on delayed recall -- that is, retention after a period of time as a function of some experimental manipulation. The researcher would like to be able to demonstrate a difference that could unambiguously be attributed to the manipulation; in other words, a difference "unconfounded" or clouded by other "irrelevant" variables that may have influenced the observed difference.

Consider Figure 6, which presents some hypothetical results from a typical retention experiment, with two groups of subjects (experimental and control) and measures of performance collected at two times ("Pre Test" and "Retention Test"). The treatment -- the manipulation that distinguishes the two groups -- could be anything hypothesized to affect retention. The hypothetical results show that the groups differ to a small extent on the pre-test and to a greater extent on the retention test. The issue is how to analyze these results so as to obtain the clearest interpretation of the effects of the treatment.

If we were primarily interested in retention score differences, we would be concerned about the possible confounding effects of pre-test score differences between the groups. In this type of situation, three things could be done statistically to aid interpretation. First, we could statistically evaluate the pre-test differences (using, for example, a t-test) and, if the difference were nonsignificant, ignore it. Second, we could statistically "remove" the variance associated with pre-test scores from the retention test score variance. This technique, the analysis of covariance, essentially subtracts out the correlation between pre-test and retention test scores from the latter. Third, we could use an analysis of variance. This technique would involve examining the interaction of groups and treatments; if the interaction were statistically significant, the interpretation of the results would be that the treatment differentially affected the groups. This "weak" conclusion could be strengthened by comparing the pre-test scores directly.

Again, if we are primarily interested in treatment effects -- that is, the retention score differences -- and the pre-test differences are irrelevant with respect to the treatment (e.g., when the treatment follows the pre-test), the analysis of covariance is the most powerful technique to use. It "sharpens up" the retention score comparison by eliminating the variance in the scores due to some non-treatment differences.

In many cases, the researcher is interested in more than the retention score differences; in many experiments, ANCOVA is an inappropriate technique. It is arguable that in the experiments we examined, "immediate posttest score" is relevant in the sense discussed above. More technically, the choice of which of these two analyses to use depends on whether the research question
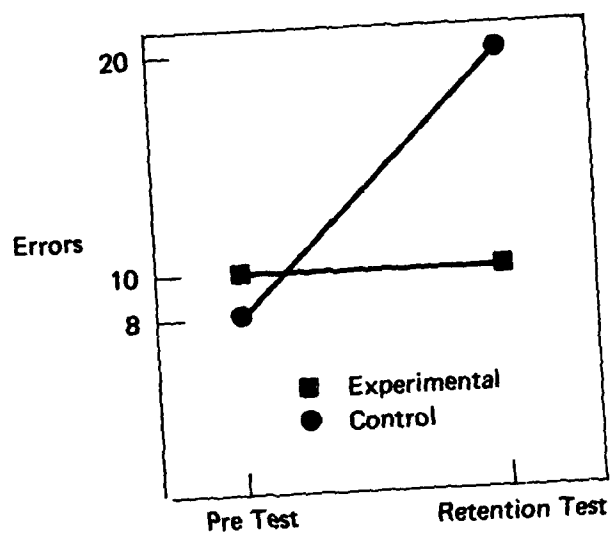
FIGURE 6. Hypothetical Results of a Retention Experiment.

being addressed focuses on performance averaged across all retention intervals (including pre-test scores) or on delayed retention only. For the present review, the choice of analysis was moot in that our ANCOVA reanalyses in practically all cases supported ANOVA interpretations.

Form of the decay function. Decay is rapid over the first hours and days after completion of training and slows as time passes. Thus, use of points on one part of the range of relevant retention intervals will provide estimates of retention that cannot be interpolated or extrapolated to other intervals. In one project which estimated linear regression coefficients, for example, the estimates were computed using data only after four months of retention, and as a result, the rates of loss are probably estimated as lower than similar estimates in other projects. To understand the importance of this problem, consider Figure 7. In the three parts of Figure 7, data from three parts of the same curved decay function are used to estimate a linear decay rate, and the results are quite different. In this figure, the solid line represents the (hypothetical) true retention curve, the same in all cases. The three points represent the actual data points collected; and the dotted lines represent the straight line fit to the three points in each case. Note that the straight line fits would predict different skill levels at different points in time although the true retention curve does not vary. Thus, the researcher must use caution when interpreting decay functions. Whenever possible, increasing the number of data points and/or increasing the inter-test intervals would be good experimental procedures.

SUMMARY AND CONCLUSIONS

As mentioned above, this report is part of a current ARI project being conducted by AIR. Our goal is to produce a convenient, practical method that will help unit Commanders and training managers to determine which soldiering tasks can be maintained at a high level of proficiency with little or no practice, and which tasks will require additional training. The ultimate aim is to enable Army Commanders and training managers to make informed decisions about the kind of training to conduct, and when, in order to ensure operational effectiveness.

The research that has been done on the acquisition and retention of tasks and specific skills that underlie task performance is critical for our needs. While the experimental work done in laboratories and reported in the general literature is helpful, the bulk of this research lacks the external validity to be generalized to the complex conditions associated with performing military tasks in actual Army settings. The ARI research has largely overcome this deficiency by examining the learning and
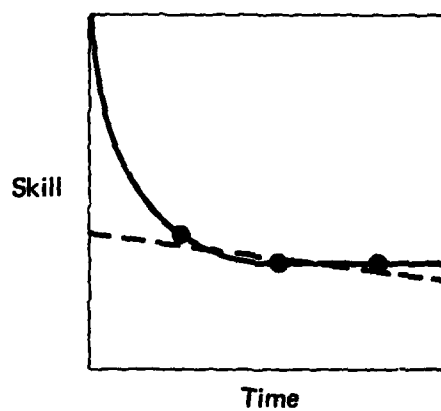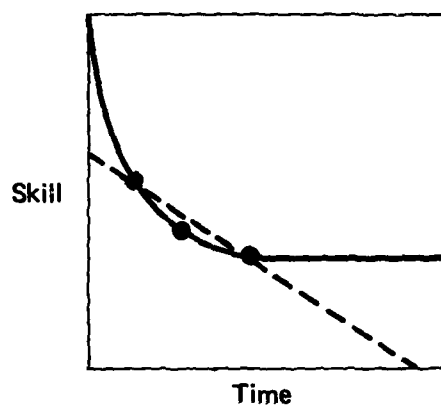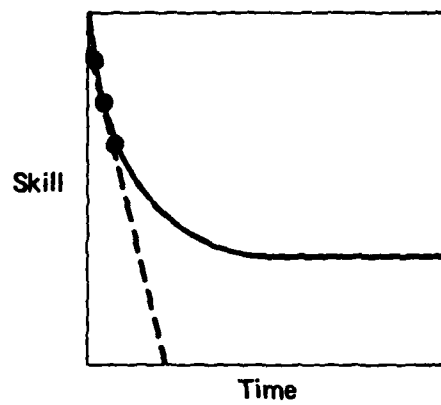
FIGURE 7. Linear Regression Estimates Based on Different Regions of the Decay Curve.

forgetting of Army tasks in Army units and schools. As such, these projects are the proper starting points for our own research.

Considered as a set, the ARI efforts have identified and investigated the key variables. They have examined the effects of individual soldier abilities, amount of training, different training strategies, job environment, task characteristics, and the retention interval of Army task acquisition and retention. While the ARI efforts are not conclusive, they have given us insight for how to structure our own research. Moreover, secondary analyses of their data will yield additional information that will help us to meet the objectives of our research.

The value of the ARI-supported research reviewed in this paper are twofold: (1) They demonstrate the research approaches, experimental designs, and measurement techniques that will generate the information needed to develop the means for predicting the loss of proficiency on Army tasks; and (2) they illustrate the need for sufficient support so that the research conducted in the field will yield results that can be applied Army-wide.

ANNOTATED BIBLIOGRAPHY OF REFERENCED PROJECTS

Dressel, J.D. Mnemonically enhanced training. Working
Paper. Alexandria, VA: ARI, February 1980.

Summary. The purpose of this project was to determine the in-
fluence of a mnemonic (i.e., memory-aiding) technique upon the
acquisition and retention of a procedural task: installation of
an M14 anti-personnel blast mine. Subjects learned the task
either with or without the use of the mnemonic technique and were
tested for task retention one month later. Results indicated
that performance was not influenced either by method of training
or by retention interval.

Training considerations. Since the task used was highly
proceduralized, it was hypothesized that a verbal memory aid
could be used to cue the sequence of steps. The aid took the
form of an "acrostic": in this case, a sentence wherein each let-
ter corresponded to a task operation and where the sentence
preserved the correct order of operations. This technique was
compared to a control condition where subjects spent an equiv-
alent amount of time studying the task, but without the use of
the mnemonic. This training manipulation was not effective in
improving recall.

Individual differences. Subjects were 51 soldiers engaged in One
Station Unit Training. General-technical (GT) score⁻ were ob-
tained for all soldiers, and used as a performance correlate.
The pattern of correlations obtained suggested that the mnemonic
technique may have increased the difficulty of the task.

Task considerations. The task used was installing the M14 an-
tipersonnel mine, a task in the MOS 12B Soldier's Manual. This
task had 18 discrete steps, which must be performed in the cor-
rect order.

Retention considerations. The retention interval used was one
month. Soldiers did not practice the task during this interval.
There was essentially no retention loss across this interval for
any of the soldiers.

Goldberg, S.L., Drillings, M., and Dressel, J.D. Mastery
training: Effect on skill retention. Alexandria, VA: ARI,
unpublished.

Summary. The purpose of this project was to determine the ef-
fects of mastery training and length of retention interval on the
retention of a procedural skill. The task used was to boresight
and zero the main gun of the M60A1 tank. Soldiers were trained
to either a standard performance level of one correct performance
or to a mastery level of three consecutive correct performances.

Retention was tested after either one or five weeks. The results were that both amount of training and length of retention interval affected performance, with soldiers performing better after the shorter interval or after more extensive training.

Training considerations. The major variable in this project was a manipulation of the level of initial training. Soldiers were trained to a criterion of either one or three consecutive errorless performances of the task. Results showed a significant improvement in retention for the more highly-trained group. However, their superiority disappeared after the first relearning trial. Also, despite mastery training, "absolute" level of skill was still low.

Individual differences. Participants in this project were 42 Armor Crewmen; all crew positions were represented. Thus, subjects differed with respect to Army experience. Results indicated that this experience affected acquisition performance in the expected direction. Also, mental category data for most of the sample was obtained; however, analyses of these data were not conclusive.

Task considerations. Two tasks were used in this effort. The first was to boresight the M60A1 tank, which consisted of 11 steps. The second was to zero the M60A1 main gun, which consisted of 16 steps. Boresighting was always performed before zeroing. Results for both acquisition and retention were examined separately for each task. Boresighting was easier to acquire; however, mastery training did not significantly improve retention. On the other hand, mastery training did significantly improve retention of zeroing.

The authors also examined individual error rates for each step of the tasks during both acquisition and retention. It was found that steps that were difficult to acquire were also difficult to retain. Furthermore, these difficult steps tended to be those that were most memory-dependent and those that were either not highly related to the preceding steps or not suggested by the hardware.

Retention considerations. Two retention intervals were used: one week and five weeks. Results indicated that retention differed for these intervals in the expected direction. This held true for both tasks and for both types of training.

Hagman, J.D. Typewriting: Retention and relearning. Research Report 1211. Alexandria, VA: ARI, June 1979.

Summary. The purpose of this project was to evaluate loss of straight-copy typewriting skill between Advanced Individual Training (AIT) graduation and unit duty, and to determine the effect of refresher training on recovery of proficiency.

Performance was measured on three occasions: at end-of-course testing, at the start of unit duty, and after a practice session. Proficiency dropped substantially between AIT and unit duty; practice reinstated some of the lost skill.

Training considerations. Baseline proficiency was measured directly at the end of AIT. Soldiers varied in skill, but all reached a minimum of 25 net words per minute.

Individual differences. Subjects were 38 Administrative Specialists, 71L MOS, and who were initially tested at the completion of AIT. Although soldiers differed in skill at initial testing, there were no analyses of individual differences.

Task considerations. The task for all tests was straight-copy typewriting. Alternative forms of the test were equated for difficulty.

Retention considerations. The retention interval varied between 14 and 38 days of no intervening practice. The inclusion of both an unpracticed test and a test after a practice period enabled the author to analyze the effects of refresher training directly; results indicated that this refresher training restored some of the proficiency lost over the no-practice interval. The restoration of skill was primarily in terms of speed; error rate was not affected by the refresher practice.

Hagman, J.D. Effects of presentation- and test-trial training on motor acquisition and retention. Technical Report 431. Alexandria, VA: ARI, January 1980.

Hagman, J. D. Effects of presentation- and test-trial training on acquisition and retention of movement end-location. Technical Report 492. Alexandria, VA: ARI, November 1980.

Summary. The purpose of these two projects was to examine the relative effects of three different training methods on the acquisition and retention of a positioning motor task. These two projects will be discussed together as they were essentially equivalent in all aspects of the design and procedures; the primary difference was that in the first, subjects were instructed to learn and remember movement distance, while in the second they were instructed to learn and remember movement end position. The different training methods were different combinations of presentation trials (where performance was guided) and test trials. In both cases, it was found that an emphasis on presentation trials improved acquisition performance, while emphasis on test trials improved long-term retention.

Training considerations. The primary experimental manipulation in both efforts was the sequence of trial types used in initial training. Groups of subjects received either alternating trials

of presentation trials (i.e., movement to a mechanical stop) and test trials (i.e., without the stop), mostly presentation trials, or mostly test trials. In the first project, the alternating trials and presentation trials groups had smaller errors at the end of acquisition and for the short-term (3 minute) retention test, while the group which emphasized test trials had better long-term (24 hours) retention. This pattern of results was replicated in the second study.

Individual differences. Both efforts used 45 subjects, obtained from the ARI staff. Individual differences were not examined.

Task considerations. The task in both efforts was a simple horizontal movement of a metal slide along a rod. Movement distance was 250 mm in the first project; in the second, movement distance varied while the end position was held constant.

Retention considerations. Two retention intervals were examined: short-term (3 minutes) and long-term (24 hours). As mentioned above, these intervals had different effects for the different training groups.

Hagman, J. D. Effects of training task repetition on retention and transfer of maintenance skill. Research Report 1271. Alexandria, VA: ARI, May 1980.

Summary. The purpose of this project was to determine how training task repetitions affect retention of maintenance skill effectiveness. Soldiers were trained and tested on the task of testing charging system output using the 500A Sun Test Stand. Soldiers were given from zero to four repetitions during training. Testing included both a retention test and a transfer test (i.e., the same task on a slightly different charging system). Results indicated that retention improved with task repetition and deteriorated over the retention interval (14 days). Retention performance was maximized after three repetitions, with no added benefit for the fourth repetition. Transfer was not differentially affected by task repetition.

Training considerations. The major experimental manipulation was the number of task repetitions performed during initial training. Different groups of soldiers practiced the task from one to four times (a control group received only familiarization training). This manipulation affected retention performance: all experimental groups retained more than the control group; furthermore, the groups who repeated the task three or four times retained more than the groups who repeated the task one or two times. For the transfer test, all experimental groups were superior to the no-practice group, but did not differ among themselves.

Individual differences. Subjects were 60 student Fuel and Electrical Repairmen (63G MOS). Subject differences were not examined in this study.

36

Task considerations. The task used in this effort, testing charging system output using the 500A Sun Test Stand, contained five segments: setting the test stand switches and controls to base positions (28 steps), attaching charging system cables to the test stand (2 steps), setting test stand switches to appropriate positions for testing (3 steps), testing electrical output (11 steps), and shutting down the test stand (8 steps). Each of these segments had an explicit performance aid, which soldiers used for both training and testing purposes. The individual segments were examined for errors. It was found that most errors occurred in the testing segment. The author argued that this segment involved a greater load on memory, possibly because the performance aid did not provide sufficient detail.

Retention considerations. The retention interval used in this study was 14 days of no practice. This interval resulted in substantial performance decrement in the expected direction. This effect did not interact with the number of repetitions.

Hagman, J. D. Effects of training schedule and equipment variety on retention and transfer of maintenance skill. Research Report 1309. Alexandria, VA: ARI, November 1980.

Summary. This effort was highly similar to the above (Hagman, May 1980), in that it used the same task and had the same retention/transfer design and retention interval. Again, soldiers were trained and tested on the task of testing charging system electrical output using the 500A Sun Test Stand. In this case, they were trained under either massed or spaced practice conditions and with or without equipment variety (i.e., three different systems). Testing included a retention test and a transfer test. Results indicated that retention performance was better after spaced practice, as was transfer performance. Equipment variety did not influence retention, but did improve transfer performance when paired with spaced practice.

Training considerations. This effort examined two important issues regarding initial training. The first was a comparison between spaced practice (three sessions, each separated by a one-day rest period) and massed practice (three successive repetitions). The second was a comparison between soldiers who trained on only one system and those who trained on three different systems. Results indicated that there were no differences among groups immediately after training; however, spaced practice groups retained more than the massed practice groups. They also demonstrated faster transfer to new equipment. The groups trained on three different systems and those trained on one system did not differ on the retention task; however, the former groups showed more rapid transfer.

Individual differences. Subjects were 60 AIT students, MOS 63G (Fuel and Electrical Repairers). Differences in skill among soldiers were not examined.

37

Task considerations. For a description of the task, see the above effort (Hagman, May 1980). Results of an examination of individual task segments were similar: the "testing" segment was again the most difficult to retain.

Performance on a transfer task was also examined. It was essentially the same task, but performed using different equipment. Results showed that superior transfer was related to spaced practice. Also, equipment variety, when combined with spaced practice, produced the best transfer.

Retention considerations. Again, as in the above case, performance was affected by the retention interval in the expected direction. This interval resulted in substantial performance decrement for the massed practice group.

Hagman, J.D., and Schendel, J.D. Effects of refresher training on job-task typewriting performance. Technical Report 410. Alexandria, VA: ARI, October 1979.

Summary. The purposes of this project were: to determine the effect of refresher training on job-task typewriting performance; to compare the relative effectiveness of two refresher training methods on job-task performance; and to examine the relationship between straight-copy and job-task typewriting performance. Two different training methods were used, one emphasizing study, the other emphasizing recall. Major findings were that both types of training methods improved job-task typewriting skill, the two methods producing equivalent improvement; and that job-task typing performance could not be predicted from straight-copy performance.

Training considerations. This was essentially a training study, rather than a retention study. The key feature was that one of the two training methods involved a "training by recall" procedure, as opposed to a method which emphasized study of on-hand materials. Initial level of proficiency was measured directly as first-session performance.

Individual differences. Subjects were 30 Administrative Specialists, MOS 71L; all were working in their MOS. Soldiers differed with regard to their initial level of proficiency; however, these differences were not separately analyzed.

Task considerations. The job-task typewriting task required soldiers to reformat and produce examples of a Military Letter, Memorandum, Endorsement, and Disposition Form. Depending upon the experimental condition, soldiers were allowed to use examples of correctly formatted materials from a standard military reference. A straight-copy typewriting task was also used; this task had alternative forms, equated for difficulty.

Retention considerations. The "retention" manipulation was to have half the soldiers use the reference materials for all trials and to have the other half try to recall the correct formats on half of the trials. This manipulation did not differentially affect job-task performance.

Holmgren, J.E., Hilligoss, R.E., Swezey, R.W., and Eakins, R.C. Training effectiveness and retention of training extension course (TEC) instruction in the combat arms. Research Report 1208. Alexandria, VA: ARI, April 1979.

Summary. The purpose of this project was to determine the training effectiveness and retention of TEC instruction relative to conventional classroom instruction. TEC lessons in five subject areas were evaluated. Results indicated that, averaged across the five areas, the TEC trained soldiers performed better than the conventionally trained soldiers on both the initial and retention tests.

Training considerations. The major variable employed was the comparison between the TEC instructional package and conventional classroom instruction. In addition to the audio-visual packages, the TEC lessons also include a Lesson Administrative Instruction (LAI) test. This project also evaluated the possible effect of these tests on retention. Thus, there were five training groups: TEC instruction with LAI tests; TEC instruction only (i.e., no LAI tests); conventional instruction with the LAI tests; conventional instruction without the tests; and a control group, which received neither instructions nor tests. Separate performance tests for each of the five subject areas were developed. These tests were administered twice: once immediately after training, and once from seven to twelve weeks later. Results indicated that the performance of the TEC trained soldiers (Active Army) was superior to the performance of the conventionally trained soldiers on both the initial and the retention tests for four of the five subject areas. Also, for National Guard soldiers, the TEC trained groups were superior to the conventionally trained ones on two areas in the initial test and one area in the retention test. There were no differences in performance between groups administered or not administered the LAI tests.

Individual differences. The Active Army segment of the experiment consisted of 630 soldiers from ten Combat Arms MOS. The National Guard segment consisted of 539 soldiers, again from various MOS. GT scores were collected for some of the soldiers. These scores were positively correlated with performance test scores for the Active Army soldiers.

Task considerations. Five TEC lesson series were evaluated. These five included one each for the four major Combat Arms:

- Field Artillery - Use of the Gunner's Quadrant
- Air Defense Artillery - Use of the TADDS system
- Armor - M551 Tank Target Engagement simulation
- Infantry - Use of the Squad Radio

A fifth lesson set on firing and zeroing the M60 machinegun was administered to all soldiers. Each of the performance tests developed for the experiment was constructed to be hands-on and was scored in terms of both tasks passed and individual steps passed. These subject areas differed on retention and as a function of whether the soldiers were Active Army or National Guard. However, these differences were not analyzed independently from the TEC-conventional training comparisons.

Retention considerations. The retention intervals used were 8-9 weeks for the Active Army soldiers and 7-12 weeks for the National Guard soldiers. Performance declined for both groups across all tasks. The authors also collected job environment information for all soldiers; this information was used to discard soldiers who had practiced any task during the retention interval.

McCluskey, M.R., Hiller, J.H., Bloom, R.D., and Whitmarsh, P.J. Skill decay of sixteen common tasks for MOS 11B and 11C. Final Report. Alexandria, VA: ARI, November 1978 (Draft).

Summary. The purpose of this project was to develop skill decay curves for sixteen common Soldier's Manual tasks for MOS 11B and 11C. Following a 2-4 week training period, all soldiers were given a performance posttest. Each soldier was then requested to return for one performance retention test at one of three retest intervals. The results of this project did not produce reliable skill decay curves.

Although this project did not produce sufficiently reliable data to be reported upon here, one aspect of the method bears discussion. The authors carefully defined two task characteristics hypothesized to relate to retention and selected tasks according to these characteristics. The two dimensions were: (1) task learning difficulty, and (2) the amount of performance guidance cues available which function either as response prompts or as feedback. For this latter dimension, the authors surmised that if task guidance is high, then the need to rely on "sheer memory" is minimized. They were able to sort the sixteen selected tasks into four categories formed by the combinations of these two dimensions.

Osborn, W.C., Campbell, C.H., and Harris, J.H. The retention of tank crewmen skills. Research Report 1234. Alexandria, VA: ARI, December 1979.

Summary. The purpose of this project was to examine the retention of armor crewmen skills from the time of training through the early months of job assignment. Performance tests of job tasks were administered to soldiers in two Armor MOS at the end of formal training and again four to eight months on the job. Results indicated that proficiency overall remained unchanged from school to the field, but when examined by task category it was found to decline for tasks common to all crewmen and to increase for tasks specific to a crew position. No systematic changes in proficiency occurred as a function of time since training or relevance of job assignment.

Training considerations. There was no manipulation of the training provided to the subjects. Initial training was Armor OSUT. The tasks used in this effort were checked against the OSUT program of instruction to insure that all tested tasks were included in training. The mean percent of tasks passed was 42.6 for drivers, 47.6 for loaders, and 22.6 for gunners. The mean percent of a sample of common soldier tasks passed by these same groups was 89.1, 84.9, and 84.9. These scores were obtained upon completion of OSUT. In general, results showed no performance decrement over time (and, in fact, a slight increase) for the job-related tasks, and a slight decrement for the common soldier tasks.

Individual differences. Subjects were 32 drivers (19F MOS) and 64 gunner/loaders (19E MOS). These numbers are only those soldiers who completed the entire test/retest cycle; initially, several more subjects were tested. Subject attrition was primarily due to administrative reasons, unrelated to test scores.

Task considerations. A total of 65 tasks were studied (for a complete list, see Table 3 of the present report). These included:

- 8 Driver tasks, 2-16 steps, no subtasks
- 27 Loader tasks, 1-18 steps, some with subtasks
- 16 Gunner tasks, 1-18 steps, some with subtasks
- 12 Common Soldier Tasks, 3-11 steps, no subtasks
- 2 Common Crewmen Tasks, 3-5 steps, no subtasks

The general trend of the results was that the job-specific tasks showed no retention loss, while common task performance declined over time.

The authors also examined the most frequently failed task elements. They hypothesized that these elements fall into two types: those that are the most difficult or skilled aspect of the task, and those in which the relevance of the task element is

questionable or unclear to the crewman. This latter category could include, for example, safety precautions that the crewman does not consider necessary in the test situation.

Retention considerations. The time interval between OSUT completion and early unit assignment varied between four and eight months. Analyses were conducted for intervals less than five months, 5-6, 6-7, or greater than seven months. These intervals did not differ in the pattern of retention losses (or gains) observed.

The authors also examined certain job-environment factors, such as whether the soldier was assigned to his correct duty position, time in duty position, and training received on the job. However, these data were too sparse and inconsistent to evaluate.

Schendel, J. D., and Hagman, J. D. On sustaining procedural skills over prolonged retention intervals. Research Report 1298. Alexandria, VA: ARI, July 1980.

Summary. The purpose of this project was to determine if long-term retention of procedural skills depends upon how periodic refresher training sessions are scheduled, and if soldiers can estimate in advance of retention testing how much training they require to regain proficiency. Soldiers were trained to disassemble and assemble the M60 machine gun under one of three conditions: until a criterion of one errorless performance was reached; extended practice (100% overtraining beyond one errorless trial); and "spaced sessions", where the 100% overtraining was provided midway through the 8-week retention interval. Key results were that soldiers did not forget substantially more over the 8-week interval than over the 4-week interval; both overlearning groups retained more and relearned faster than the normally-trained group, but did not differ from each other; and that soldiers could predict how much refresher training they would need to regain proficiency.

Training considerations. A major variable in this effort was the type and amount of initial training received by the soldiers. Three different methods were used: the first required soldiers to practice the task until a criterion of one errorless trial was performed; the second group received 100% overtraining -- they received additional trials equal to the number of trials necessary to reach errorless performance; and the third group received the same 100% overtraining, but four weeks after the initial learning. The overtraining for this third group could be considered equivalent to refresher training. Results from the retention tests showed that the overtraining (whenever it was administered) significantly improved both retention and speed of relearning.

Individual differences. Subjects for this effort were 38 Army reservists. Treatment groups were matched by age, sex, and Army experience. Data were collected regarding soldiers' estimates of their performance capabilities for the test task. Results indicated that, as a group, they could accurately estimate the number of retraining trials necessary to regain previous proficiency.

Task considerations. The task employed was to disassemble and assemble the M60 machinegun. As used in this situation, this task was slightly modified and simplified. Soldiers required about 30 minutes to reach criterion performance (approximately two trials). No analyses of task components were reported.

Retention considerations. Task retention was examined at two intervals (four and eight weeks), and between the overlearning groups and the control (no overtraining) group. A comparison between the refresher training group's performance after four weeks (and before refresher training) and the control group after eight weeks showed no differences; that is, soldiers forgot as much after four weeks as after eight. The other comparison, between the three groups after eight weeks, showed that retraining after four weeks was equivalent to providing overtraining initially: the two overtrained groups were equivalent and substantially better than the control group.

Shields, J.L., Goldberg, S.L., and Dressel, J.D. Retention of basic soldiering skills. Research Report 1225. Alexandria, VA: ARI, September 1979.

Summary. The purposes of this project were to evaluate soldiers' retention of basic skills learned in initial training and to determine how task factors affect skill retention. It was found that tasks varied in the rate at which the percent "GO" declined since training. Differences were attributed to the number of task steps, the presence or absence of subtasks, and the order of original training. Also, the steps that were forgotten tended to be those that were not suggested by the previous sequence of steps or by the equipment.

Training considerations. Initial levels of task proficiency were not obtained for the soldiers. It was assumed that they could or did pass all tasks at the completion of training. A brief "coaching" session, in which soldiers saw a demonstration of task performance and had a short practice period, was included as an experimental treatment. This treatment had no systematic effect on the retention results. Any soldier who reported having practiced a task in the interval between graduation and testing was not included in analyses of retention. Two task training variables were included in the analyses: the serial order in which the tasks were trained, and the number of repetitions each task received during training. The former was shown to be a predictor of retention loss for certain tasks.

Individual differences. The subjects were 523 soldiers,
including 182 who were taking an end-of-course test from Basic
Training or One Station Unit Training, and 341 soldiers with
various MOSs (primarily 13B, but also including other field ar-
tillery MOSs). Mental Category data were obtained for the sol-
diers, but were not further analyzed.

Task considerations. Tasks studied were:

>   Cardiopulmonary Resuscitation
>   Stop Bleeding
>   Challenge and Password: One Man Approaches
>   Challenge and Password: Group Approaches
>   Report Enemy Information: Size/Activity/Location
>   Report Enemy Information: Unit/Time/Equipment
>   Don Protective Mask
>   Individual Defensive Position - Outline
>   Individual Defensive Position - Describe
>   M60 Machine Gun - Load/Fire
>   M60 Machine Gun - Reduce Stoppage
>   M60 Machine Gun - Unload/Clear
>   M203 Grenade Launcher - Disassemble/Assemble
>   M203 Grenade Launcher - Load/Fire
>   M203 Grenade Launcher - Reduce Stoppage
>   M203 Grenade Launcher - Clear
>   M72 Light Anti-tank Weapon (LAW) - Inspect/Fire
>   M72 LAW - Restore
>   M16 Rifle - Disassemble/Assemble
>   Communications Check

These tasks varied along several dimensions that the authors ex-
amined, including the number of steps in the task, whether tasks
had safety procedure steps, and whether tasks could be broken
into subtasks. These dimensions were shown to be predictors of
retention loss.

Retention considerations. The basic design was a cross-sectional
time series, where the retention interval varied between four and
twelve months. Slopes were calculated which reflected the loss
of performance over time for each of the tasks.

>   Shields, J.L., Joyce, R.P., and VanWert, J.R. Chaparral
>   skill retention. Research Report 1205. Alexandria, VA:
>   ARI, March 1979.

Summary. The purpose of this project was to evaluate retention
of Chaparral skills and to determine the most effective schedule
of refresher training. Soldiers were tested immediately after
AIT on several Chaparral tasks, retested upon arrival in their
battalions, and again retested four months later. Comparison
groups were tested after a one- or two-month retention interval.
Results showed that, in general, performance did not

substantially decline for most of the tasks; this was attributed to the use of job aids for task performance.

Training considerations. Perhaps the key feature of this effort was the extensive use of job and performance aids in training. Five TEC lessons, plus TMs (all "easily obtained"), were supplied. Soldiers were instructed in the use of these materials. The authors argued that these aids were responsible for the lack of retention loss observed.

Individual differences. Subjects were 71 (completed) Chaparral crewmen (16P MOS). Each subjects' AIT test score (i.e., level of initial performance ability) was used as a predictor of retention scores. It was found that these AIT scores were predictive of retention performance for all but one task.

Task considerations. Six Chaparral tasks were used:

- Pre-energizing the M-54 launch station
- Energizing the M-54 launch station
- De-energizing the M-54 launch station
- Before operations PM checks on the M730
- Installing and operating the TA-312/PT telephone set
- Emplacing, operator checks, and adjusting target alert data display set AN/GSQ-137

These tasks differed slightly in retention scores and their predictability from a soldier's AIT scores. The authors did not discuss any hypotheses for these differences.

Retention considerations. The time intervals used -- one, two, and four months -- did not result in significant performance losses for any of the tasks. Again, the authors argued that the job and performance aids used substantially reduced the effect of the retention interval.

Sullivan, D.J., Casey, R.J., and Hebein, J.M. Acquisition and retention of cognitive versus perceptually oriented training materials. Technical Report. Alexandria, VA: ARI, October 1978.

Summary. The purpose of this project was to investigate the effect of various instructional strategies on the acquisition and retention of aircraft ranging and engagement skills across individuals varying in academic aptitude. Instructional strategies tailored to different aptitudes of soldiers were developed and evaluated against task performance and retention of a simulated Redeye missile crewman task, using a civilian population. Another experiment, using military subjects, was also conducted. Analyses of the retention data were limited to an evaluation of the instructional strategies; results indicated that the use of

45

aptitude measures, if carefully chosen and validated, may be an effective and efficient approach for determining the most appropriate instruction for military trainees.

Analyses and discussion of the retention phase of the experiment were not presented in sufficient detail in the report to allow any inferences to be drawn for purposes of the present report. The reader is referred to the referenced report for further details.

Vineberg, R. A study of the retention of skills and knowledge acquired in basic training. Technical Report 75-10. Arlington, VA: ARI, June 1975.

Summary. The purpose of this study was to measure retention of basic combat skills. Thirteen subtests of the Comprehensive Performance Test (CPT) were administered to soldiers during the last week of BCT and again six weeks later during AIT. Results indicated that the probability of the average soldier passing a CPT subtest at the end of BCT was 0.81, of passing during retention testing six weeks later, 0.63, and of passing both at the end of basic training was 0.55. Depending on the measure of retention used, there was an average decrease in proficiency of from 18% to 26%. Decrements in performance varied across subtasks and Mental Categories.

Training considerations. There was no manipulation of the training provided to subjects. Training emphasized performance-based instruction through actual performance of tasks; no information on actual regimens was provided. The proportion of soldiers passing a given test ranged from 1.0 to 0.417 during the initial testing.

Individual differences. Subjects were 200 soldiers who graduated from BCT at Ft. Ord, who remained at Ft. Ord for AIT, and who provided test data both times. The subjects differed with respect to AFQT levels: 44 at Level II, 120 at Level III, and 36 at Level IV.

Task considerations. Thirteen (hands-on) tests representing seven subject areas were drawn from the Comprehensive Performance Test (SMART, 1 April 1974). They included:

- Drill and Ceremonies (D&C)
    Marching movements for the individual
    Manual of Arms executed from sling arms

- First Aid (FA)
    Apply tourniquet, dress wound, treat for shock
    Treat for burns, treat for shock

- Individual Tactical Training (ITT)
      Passage of obstacles during daylight
      Individual maneuver techniques

- Guard Duty (GD)
      Inspecting Officer
      Hour of darkness with proper or improper authorization

- M16Al Rifle (M16)
      Clearing the weapon
      Immediate action

- Chemical, Biological, and Radiological (CBR)
      First aid for a nerve agent casualty
      Reaction to nuclear burst without warning

- M60 Machine Gun (M60)
      Placing the M60 machine gun into operation and
      performing immediate action.

Soldiers had to perform each of one or more required steps in a prescribed manner and sequence in order to pass the test. For individual subtests of the CPT, the average retention loss varied between 5% (Inspecting Officer) and 40% (Clearing the M16Al rifle). Tests were sampled by level of difficulty but the effect of this variable was not analyzed.

Retention considerations. The retention interval was six weeks. A questionnaire was administered as a prelude to the second testing to determine if soldiers knew they were going to be retested and whether they had practiced or studied their SMART books. Only five did so.

The retention measure was based on the mean proportion of tasks passed initially minus the mean number passed subsequently. The two were correlated in that the tests retained their ordering across sessions.